
S-114.100 Computational Science / Laskennallinen tiede. Fall 2004.

Assignment 7. Smoothing of data and method of least squares
Chapter 7 in Lecture notes.

Due Thu 10.11.2003 (3 problems, total of 5 points)

Web page: www.lce.hut.fi/teaching/S-114.100/

computer = programming task (C / C++ / Fortran / Java)

pencil and paper = solve on paper

Problem 1. (*pencil and paper*)

The viscosity V of olive oil as a function of temperature is known to vary according to a quadratic law $V = a + bT + cT^2$. Using the principle of least squares, find the best values of a , b and c for the following table of data:

T	295	298	304	308	318	328	337
V	73.6	63.3	49.8	41.7	28.9	31.0	40.5

Problem 2. (*pencil and paper / Matlab*)

The following set of data

x	0	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3.0
y	1.40	1.45	2.42	3.05	3.63	4.63	6.46	8.26	10.55	13.95	18.02

follows a relationship $y = ae^{bx}$. In order to find the appropriate values of a and b , apply the least-squares principle after linearizing the problem. You are allowed to use mathematical software for finding the least-squares line (e.g. the Matlab command `polyfit` can be used for the line fit).

What are the parameters for the line fit? What are the corresponding values of a and b ? Show a plot of the resulting nonlinear fit together with the data points (x_i, y_i) . Is the solution entirely correct?

Problem 3. (*computer*) (3 points)

(a) Write a program which uses the polynomial regression algorithm to produce a polynomial fit of a given set of data. The output of the polynomial regression procedure should be α_i , β_i , σ_i^2 and c_i for $0 \leq i \leq n$ where n is determined by the condition $\sigma_{n-1}^2 > \sigma_n^2 \approx \sigma_{n+1}^2$.

Apply the program for the following set of data:

x	y
-9.70	3.76
-7.30	1.78
-5.40	1.52
-5.00	1.31
-3.01	0.31
-2.13	0.23
-1.20	0.45
-0.56	0.29
0.00	0.00
1.20	0.45
4.50	0.28
6.70	2.12
9.90	3.91
10.00	3.47
12.30	5.59

Evaluate the polynomial at 100 equidistant points in the interval $[-9.7, 10.0]$. Plot the polynomial and the data points. What is the degree of the fitted polynomial?

(b) Try to use your polynomial regression program on the following *world population* data (year vs. population in billions):

x	y
1000	0.340
1650	0.545
1800	0.907
1900	1.61
1950	2.51
1960	3.15
1970	3.65
1980	4.20
1990	5.30

What happens? Why? Under what conditions is it meaningful to use polynomial regression?